DOCUMENT RESUME

ED 184 130 CS 205 440

AUTHOR Hudson, Sally; Veal, Ramon
TITLE Assessing Writing Competence through Writing Samples. Studies in Language Education Report No. 36.
INSTITUTION Georgia Univ., Athens. Dept. of Language Education.
PUB DATE Oct 79
NOTE 13p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Methods; *Minimum Competency Testing; Secondary Education; *Writing (Composition); *Writing Skills
IDENTIFIERS *Writing Evaluation

ABSTRACT

This report discusses a plan to help Georgia school systems develop their own evaluation instruments for student writing. The report indicates that the use of actual writing samples in assessing writing skills can be valid, reliable, and economical. It includes attachments with sample writing scores from a school, the purposes and statistics that are relevant to various levels of the educational process, a description of the development and use of a sample writing instrument, a sample test, a sample rating sheet, correlations between two raters of tests, and a description of expenditures involved in administering standardized tests and a writing sample test. (TJ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED184130

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ASSESSING WRITING COMPETENCE THROUGH WRITING SAMPLES

by

Sally Hudson

and

Ramon Veal

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sally Hudson
Ramon Veal

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Studies in Language Education, Report No. 36
Department of Language Education, The University of Georgia
Athens, Georgia

October 1979

CS205440

ASSESSING WRITING COMPETENCE THROUGH WRITING SAMPLES

Sally Hudson and Ramon Veal

Testing writing skills through actual writing has been the plan and practice of many English teachers for more than 100 years. (See English Journal for April 1979.) With greatly increasing student populations and the development of objective standardized tests, the writing sample as a means to assess composition skill has declined in use. After all, objective tests are fairly reliable; they do correlate with other measures of writing; and they are economical. As with many other instructional practices, the more efficient method, in this case the standardized test, won out--at least in many day to day practical situations for evaluating or predicting writing skill. Thus, the following conclusions could be summarized in 1971.*

1. Only samples of writing actually lead to direct measures of writing ability.

2. Measures based on writing samples are, however, less reliable than measures obtained with objective tests.

3. Objective tests can measure knowledge and skills related to writing; they do it reliably and they can be validated against such criteria as course grades, teachers' evaluations, and reliable ratings of essays.

4. The best estimate of composition ability would be a combination of less direct but more reliable objective items with a more direct but less reliable writing sample; however, the time and expense involved in administering and scoring actual writing samples lead most test makers and users to prefer the shorter, more reliable objective tests.

Still, researchers, test developers, and some teaching institutions continued to develop and use direct methods of evaluating writing, writing samples. Even with the high costs and difficulties involved in using any-

*Veal, L. R. and Ellis, W. G. "College English Placement Tests," Journal of Educational Measurement, Volume 8, No. 3, Fall, 1971.

thing other than machine scoreable paper-and-pencil tests, the National Assessment of Educational Progress (NAEP) has been able to assess actual writing performance, and the College Entrance Examination Board (CEEB) has returned the essay to its SAT testing format. And state competency programs, like the one in Georgia, usually list a specific competency such as "The student writes legible, appropriate personal and career communications on a functional level."

The Georgia Competency Based Education plan has provided ten pilot school systems the special opportunity to develop their own evaluation instruments (tests, checklists, observations, writing sample, etc.). Exercising this option, one pilot system, Gwinnett County, developed procedures for obtaining, rating/judging, and reporting writing scores coming from writing samples. On a trial basis last May, Gwinnett County collected two pieces of writing (short essay and business letter) from a sample of 800 eighth graders and 800 tenth graders. They represented six different schools.

Following procedures clearly identified by research and experience,** at least two trained raters (Gwinnett County English teachers) used two rating scales to rate the two samples from each student on content and mechanics (see Attachment A). Repeated reliability checks yielded correlation coefficients in the .80's (see Attachment B). These empirical reliability estimates plus content (topics, criteria, etc.) validity checks

---

**See NCTE publications: Evaluating Writing by Charles Cooper and Lee Odell and Measuring Growth in English by Paul Diederich. Also note the CEEB booklet, "The English Composition Test with Essay" and, for short reviews, see NCTE SLATE publications, "Evaluating Writing" and "Standards for Basic Skills: Writing Programs" (February and April '79).

with English teachers illustrate the high level of confidence one can have in these procedures. Costs were closely watched as well; and for this trial, the cost per student (collecting, scoring, and reporting) was slightly less than $3.00, close to the price of other tests (see Attachment C). Thus, this pilot study shows that the use of actual writing samples in assessing writing skills can be practical as well as desirable--that is, it can be valid, reliable, and economical.

In this pilot study, Gwinnett County teachers and administrators received individual student scores (content, mechanics, and total) for two types of writing (essay and letter). They also received average school and grade level scores for essay and letter along with indications of rater reliability as already noted. Reporting included results like those shown in the sample which follows.

In addition to having valid, reliable, and economical scores for evaluation via rated writing samples, these procedures have already influenced English instruction in Gwinnett County. Observations by system-level curriculum consultants indicate that writing instruction and practice now includes more emphasis on letter writing. The use of set criteria for evaluating other kinds of writing has increased. And several schools are obtaining and evaluating writing samples as pre and post tests in program evaluation.

If both evaluation and instruction are improved through the use of writing samples to test writing, then their use seems critical if the teaching of written composition is as important as state and system competency goals suggest.

SAMPLE INDIVIDUAL STUDENT/CLASS

REPORTING FORM

School: XY2 Writing Task: Essay (or letter)

Scores reported as content/mechanics (content: 12-60) (mechanics: 8-40) (total: 20-100)

Jones: 2nd period

High 90-100
- Allen, B. (60/38)
- Amsted, L.(57/33)
- Astmov, V.(50/40)

High Average 70-89
- Allcorn, S. (50/20)
- Ambrosio, D. (55/29)
- Bumsted, A. (47/32)
- Calhoun, S. (40/30)
- Dedalus, F. (53/18)
- Fallcorn, F. (42/38)
- Middleton, T. (45/40)

AVERAGE 50-69
- Ahcee, T. (28/34)
- Adams, D. (34/35)
- Adams, S. (25/28)
- Bell, T. (30/27)
- Cochran, S. (24/30)
- Dolmer, M. (32/30)
- Feinstein, D. (36/30)
- Guthrie, A (27/29)
- Martin, S. (36/30)
- Martinez, M. (37/31)
- Nachez, P. (40/10)
- Numbrio, P. (30/30)
- Stachey, Z (39/30)
- Stone, V. (32/34)
- Vittoria, L. (25/25)
- Williams, L. (36/32)
- Williams, B. (29/31)
- Young, H. (30/36)

LOW AVERAGE 30-49
- Abraham, A. (15/15)
- Allison, D. (20/25)
- Anniston, F. (17/19)
- Ball, S. (20/21)
- Bestwick, V. (27/22)
- McNamara, S. (15/19)
- Triplitt, L. (39/10)

School/Grade Mean = 54.2
School/Grade S.D. = 13.9

LOW 20-29
- McDuffy, B. (10/14)
- Neistrom, R. (14/15)
- Stalmers, I. (10/10)
- Walker, D. (16/10)

These writing scores serve several different users and purposes, all with roughly the same data and relatively simple statistical procedures. Possible distinctions are shown below under three headings.

USER	PURPOSE(S)	DATA/STATISTICS
Individual Student	...to discover his/her "functional level" of achievement in written composition and to show "growth" from year to year	Composite ratings for individual written assignments--content, mechanics, and total
Teacher	...to learn individual and group levels of achievement, to use these as guides to instruction, and to indicate individual or group/class growth	Individual and group composite ratings--class means for content, mechanics, and total scores
School	...to use grade and class composite ratings in program (instructional) evaluation and revision, and to evaluate individual and group performances in terms of functional levels as well as in year to year growth	As above by school, and with correlations between raters
System	As above, but with possible comparisons among schools and grades	As above, with variance estimates
State	As above, but with possible comparisons among school systems, with specific focus on cheeking "functional writing" for graduation and reliability of raters for verification of procedures	Ratings, means, and variances
Researcher	...to check and recheck both reliability and validity of ratings as well as examine writing scores in relation to fluency, errors, mode, syntax, and other features of writing	As above, with such special analyses as word counts, error counts, mode classification, and syntactic identification

ATTACHMENT A

DEVELOPMENT AND USE OF A WRITING SAMPLE INSTRUMENT

1. Appoint Ad Hoc Committee of English teachers to complete 2 - 6 below.
2. Determine criteria typically used in evaluating written composition at specified level.
3. Select topics to be used as a basis for writing, topics narrow enough to elicit an essay in 30 minutes but broad enough to reflect common student experience, not specific knowledge.*
4. Obtain (under standardized conditions of time, directions, topic, etc.) writing samples from 8th and 10th graders on the topics to be used in the actual testing.
5. Select from these papers comparison essays that represent -- according to the criteria -- high, middle, and low quality.
6. Rate this sample of papers, using the criteria already identified.
7. Administer pilot test.
8. Train teacher-raters and rate all samples
9. Summarize and analyze ratings (descriptive statistics, including rater agreement checks).
10. Develop reporting procedures and return results and analyses to teachers, schools, and students.

---

*Short, specific tasks (like a business letter) can be used as well.

ATTACHMENT A

NAME:_____

GRADE:_____

SCHOOL:_____

DATE:_____

Atlanta Journal, May 10, 1979

WANT ADS

Briarwood Swim and Racket Club needs part - time helpers to do the following jobs:

Check passes,
work in concession stand,
clean-up grounds,
life guard,
teach swimming

If interested write to: Mr. John Jones, Supervisor, Briarwood Swim and Racket Club, 4110 Ivy Lane, Atlanta, Georgia 30248. Include your qualifications for the job desired, day and times available, and your name, address, and phone number.

Pretend you are T.J. Allen, age fifteen. You live at 610 Poplar Lane, Atlanta, Georgia 30248, and your phone number is 728-0000. You have read the ad above and are interested in one of the jobs. Using the attached paper, write a letter and address the envelope to apply for any one of the jobs. You may make up information or use your own experience.

ESSAY

Choose a thing or a person you would like to be.

You may want to be a movie star, a sports hero,

a feather, anything.

Explain what you will look like, what you will do

and why you want to be that thing or person.

ATTACHMENT A

READER#_____ PAPER#_____

	Low	Low Average	Average	High Average	High	
Ideas	2	4	6	8	10	
Organization	2	4	6	8	10	
Wording	1	2	3	4	5	
Flavor	1	2	3	4	5	_____
Usage	1	2	3	4	5	
Punctuation	1	2	3	4	5	
Spelling	1	2	3	4	5	
Handwriting	1	2	3	4	5	_____
				Sum		_____

Holistic Rating Scale for Job Application Letter

General Qualities of Content	Low		Average		High
	2	4	6	8	10
Includes sufficient qualifications & personal information Refers to specific position Uses appropriate tone (level of formality) Arranges content in effective order, excludes irrelevant information					
Conventions of Writing	1	2	3	4	5
Observes the following conventions to an appropriate degree: Effective word choice, diction Appropriate usage Effective syntax (sentence structure) aragraphing Legibility, neatness Spelling Punctuation Format - business letter					

Total score: _____

ATTACHMENT B

OVERALL CORRELATIONS BETWEEN TWO RATERS (I & II)

	CONTENT	MECHANICS	TOTAL
ESSAY	.80	.73	.87
LETTER	.73	.66	.80

Day 1

1 and 2	.90
3 and 4	.80
5 and 6	.47
7 and 0	.79
8 and 9	.93

.82

Day 2

6 and 2	.73
8 and 1	.46
0 and 3	.80
9 and 5	.84
4 and 7	.73

.74

Day 3

4 and 2	.44
1 and 7	.58
0 and 8	.82
5 and 6	.86
3 and 9	.93

.83

Day 4

2 and 7	.44
1 and 0	.91
4 and 9	.93
3 and 8	.81
5 and 6	absent

.83

During the week-long rating sessions, a reliability check was made daily. The average (via Fisher z) was .80.

ATTACHMENT C

WRITING SAMPLE EXPENDITURES

Language Arts Coordinator --11 days @ $75.00/day	$ 825.
Consultant --4 days @ $100.00/day	400.
Raters --10 for 6 days @ $45.00/day	2,700.
Clerical Help --10 days @ $35.00/day	350.
Printing	30.
Computer Service	250.
TOTAL	$ 4,555.

Cost per student (Approx. 1600) $2.85

- -

APL STANDARDIZED TEST EXPENDITURES

APL TEST (Including printing test booklet and answer sheets, use of APL Item Bank, scoring and reporting of individual and group scores.)	$ 9,000.
CBE COORDINATOR (20 days work at $90.00 per day)	2,000.
TRAVEL (To Iowa City and at various places around the state)	450.
SUBJECT CONSULTANT (3 days X 4 consultants at $85.00 per day)	1,000.
TEST ADMINISTRATION (3 middle schools and 3 high schools - Counselor's time for 2 days at $75.00 per Counselor)	450.
TOTAL	$ 12,900.

Cost per student (Approx. 1600) 8.06